DOCUMENT RESUME

ED 129 877 TM 005 679

AUTHOR Matluck, Joseph H.; Mace-Matluck, Betty
TITLE The MAT-SEA-CAL Instruments for Assessing Language Proficiency.
PUB DATE [Apr 76]
NOTE 15p.; Paper presented at the Annual Meeting of the American Educational Research Association (60th, San Francisco, California, April 19-23, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Bilingual Education; Bilingualism; *Bilingual Students; *Elementary Education; *English (Second Language); Field Studies; Grammar; Language Planning; *Language Proficiency; *Language Tests; Needs Assessment; Phonology; Reading Achievement; Scoring; *Standard Spoken Usage; Test Construction; Test Reliability; Test Validity

ABSTRACT

The Multilingual Test Development Project, recently completed in Seattle, Washington, has yielded a series of comparable oral proficiency tests in six languages designed for use with elementary school children, K-6. Extensive field-testing of the English, Spanish, Cantonese, Mandarin, Tagalog, and Ilokano tests was conducted with a research sample of several thousand students, both in the metropolitan area of Seattle and in several of the large migrant centers of eastern Washington and southern Idaho. This paper describes the project, reports on the preliminary findings, and focuses on the problems encountered and the solutions devised in attempting oral language assessment in a multilingual community. Both the findings and the research design will have implications for educational planning and program development by school districts throughout the country and for further research in these areas. (Author/BW)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

8.04

ED129877

THE MAT-SEA-CAL INSTRUMENTS FOR ASSESSING LANGUAGE PROFICIENCY

Joseph H. Matluck
The University of Texas at Austin

Betty Mace-Matluck
Seattle Public Schools

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

AERA Session 8.04
April 20, 1976

TM005 679

If we are to design adequate instructional programs in the schools of any multilingual community for children whose native language is not the dominant language of the country, we must know the level of the child's audio-lingual proficiency in both languages. To do this, we must have tests in each of these languages which accurately measure that proficiency.

Development of the Instruments. Several years ago the Seattle Public Schools became deeply concerned with this problem. Because of a series of factors involving immigration patterns and movement of populations, they were faced with a great number of children at all levels of instruction who spoke a language other than English and whose control of English was so inadequate that they could not function in the regular school program.

This has also been the special concern of the Center for Applied Linguistics in Washington, D.C., heavily involved in applied linguistics, in TESOL, and more recently in bilingual education.

The Multilingual Test Development Project, on which we have worked for the past three years under the auspices of the above-mentioned organizations, was an attempt to correct some of the ills prevalent in most oral language tests. The MAT-SEA-CAL Oral Proficiency Tests (Matluck-Matluck, 1974) in English and five other languages were developed for use with young children. The prototype is an English test, totally oral, for use in the early grades (K - 4), designed to determine the child's ability to understand and produce the distinctive characteristics of spoken English, to express known cognitive concepts and handle learning tasks in English, and to provide placement and instructional recommendations with respect to alternative instructional programs such as bilingual education, special English instruction, etc.

With the help of native informants, we then developed comparable tests with

similar objectives in Cantonese, Mandarin, Tagalog, Ilokano, and Spanish, languages which reflect the five largest of those forty non-native-English-speaking groups in attendance in the Seattle Public Schools.

Each test uses various modes of assessment: Listening Comprehension, Sentence Repetition, Structured Response, and incorporates a wide range of language features which are broken down for analysis and scoring into 1) phonology and 2) grammatical and lexical structures.

Rationale. We approached the development of the tests from the point of view of identifying some basic communication concepts that a child must handle in order to perform in a school setting, i.e. the skills of identifying, classifying, quantifying, interrogating, and negating and of showing important relationships such as spatial, case, and temporal. We then determined the grammatical manifestations of the language - English for example - the child must handle in order to perceive or to communicate these concepts. The other languages express the same concepts as does English, but each one almost always does it in different ways, i.e. with different language manifestations.

A couple of examples:

1) In order to quantify, English inflects nouns (/-s/, /-z/, /-ɨz/), but not adjectives nor verbs (except very partially in one tense only). Spanish inflects nouns, adjectives, and verbs for plural. Mandarin and Cantonese inflect none of these for plural, but use other means to pluralize nouns: numerals and certain adjectivals; Tagalog and Ilokano inflect neither nouns nor verbs for plural but do inflect adjectives and use pronouns and certain adverbs to indicate the plural subject of verbs.

2) In order to show temporality, English does it almost entirely through the verb structure, as does Spanish, but in different ways (i.e. the Spanish verb inflectional system is much more extensive). In Mandarin and Cantonese

verbs don't reflect temporality at all, and this concept is signalled by adverbs of time and some particles which change the status of a condition, occurance, or action. Tagalog and Ilokano do it partially through the verb structure (but in a different way from English and Spanish), and partially through adverbs of time..

Description of the Instruments. Each section of the test is accompanied by a master chart whose coordinates reflect the relationship between the concept and the language manifestation for each grammatical item tested.

Two of the three parts of the test in each language have a taped stimulus; the stimulus for the third (Structured Response) is supplied by the examiner. Of course, native speakers of the language being tested are used to tape, administer and score the test in each language. All of the student's oral responses are taped and all items in the test have a visual reference accompanying and reinforcing each aural stimulus, with the exception of a small "Response-to-Commands" section.

The heart of the system is, of course, the Examiner's Handbook which contains, among other things, complete instructions for administering and scoring the tests. For each section of the test there is a sentence sheet, sample visuals, a sample tape script, a vocabulary inventory, a phonology chart, and to facilitate scoring, a set of recap sheets. For each sentence of the test, there is a recap sheet containing: 1) the stimulus sentence; 2) the concept and the language manifestation involved in each item being tested in that sentence; 3) each test item extracted from the sentence and placed in the appropriate box (phonology, morphology, syntax, or vocabulary); 4) a check-box identifying those sentences which exemplify "classroom style"; and 5) at the bottom of the page, a reduced version of the visual used for that sentence.

A complete test packet contains: 1) the master chart for each section; 2) a sentence sheet containing all the test sentences of that section; 3) a recap sheet for each test item; 4) a vocabulary inventory; 5) a phonology chart; 6) a set of visuals; 7) scoring sheets; 8) audio tapes; and 9) an Examiner's Handbook.

Scoring Procedures. Scoring and diagnostic analysis are accomplished by the use of prepared charts which require the scorer to simply circle the number corresponding to the items missed. Scoring is objectivized by a system which ignores all but the particular portion of the sentence being tested. Percentage scores of each section of the test are averaged for the final composite percentage score and a line graph is supplied which allows the teacher to obtain an overall visual diagnostic profile by simply circling the number in the graph corresponding to the child's score on each concept.

The scoring procedure we designed supplies, for diagnostic and planning purposes, a profile of 1) the student's ability to handle each concept, and 2) his areas of strength and weakness. Since all of the tests are comparable in their organization and structure and in linguistic content, by administering the test in English and in the child's native language, we can compare the profiles and determine whether the child's problems are linguistic or perhaps developmental in nature.

Field Testing. The six tests were extensively field-tested in the Seattle area and in several districts in Eastern Washington in 1975 and in three Southern Idaho school districts in 1976. A total of some 5,000 tests were administered. A partial analysis of the Seattle and Eastern Washington data was completed this year. The Idaho data has been collected, compiled and is in the process of analysis. However, since this analysis is not yet completed, we will be reporting here only on the Washington data.

The tests were originally designed with the dual purpose of subjecting them to thorough field-testing and of collecting data for the Seattle Public Schools as a part of a language needs assessment.

The tests were administered to students in grades K - 6 in ten elementary schools. These schools were selected, on the one hand, because they were prospective sites for bilingual education programs, and on the other, because they contained concentrations of students whose home language was other than English.

The total sample in Seattle consisted of 609 students, in grades K - 6, 86 of whom were a control group of monolingual English speakers selected by random sampling in four of these schools. The other 523 comprised an experimental group of students identified by the teachers in each of the schools, in a Seattle Schools district survey (February, 1975), as having a home language other than English.

Tests were administered to the experimental group in English and in their home language with an interval of two weeks between tests insofar as was possible. Contamination (in the Listening Comprehension test, Part I) was minimized by attempting to administer the English test first to about one-half of the group and the home language first to the other half.

Testing time varied from about twenty-five minutes to thirty-five minutes, depending on the language being tested. Tests were usually scored by the examiner on the day after they were administered.

All examiners were native speakers of the language they administered. Each of them had at least a B.A. degree, and most were teachers in Seattle schools. Over 200 of the 609 English tests were administered by a single examiner working full time. Most of the 314 children in the largest experimental group, the Cantonese, were tested in their home language by two examiners working full time. The Spanish testing was done by one examiner, as were most of the Tagalog and

most of the Mandarin, and the Ilokano mainly by two examiners.

During the course of the field-testing, we were in constant contact with the examiners in order to elicit from them any comments or criticisms. In addition, an item analysis was completed by language group, with the dual purpose of diagnostic analysis and possible item revision.

Test data on each student in the sample was collected on prepared data sheets by school, by language group and by grade level. Scores were then tabulated and tables prepared for each of the schools by language and grade level which show mean scores, median scores and range of scores at each grade level.

Scores were re-tabulated by composite groupings and composite tables were charted which permit of multiple comparison of scores and other test data. For example, in the case of a child in the experimental group at Beacon Hill school who was tested in English and Cantonese, both his Cantonese and English scores can be compared to 1) those of his peers in his school, 2) those of the entire Cantonese group, 3) those of the monolingual English-speaking control group in his school, and 4) those of the Composite Control Group. Thus, each of the experimental groups has four composite tables at hand for comparison purposes.

Objectives. The three main objectives toward which the tests were designed were language dominance, language proficiency, and diagnostic analysis.

As to dominance, since the tests in all six languages are comparable in their construction, design, and content, a fairly objective estimate of the child's language dominance should be readily available by a comparison of his proficiency in each of his languages, as represented by his percentage score in both languages and also by an analysis of his diagnostic profiles in each of them.

We based our determination of dominance on a 10-percentage-point spread. If the child's score in Tagalog, for example, were 90% and his English score

only 50%, his Tagalog dominance would be rather obvious. If his English score were 80%, however, we would consider him balanced - 90% in Tagalog; 80% in English; with a 10-point spread. If a student's score were within a 10-point spread and both scores were over 70% (our minimum proficiency level), then we classified him as high balanced; if either of his scores was below 70%, then we considered him low balanced.

As to proficiency, a fairly accurate barometer of language proficiency in the English test was available in the form of a sizeable control group of monolingual English speakers, none of whom scored below 70%. We had no control group available of monolinguals in the home language, but once again the comparability of the tests enabled us to set this same percentage figure (70%) as a fairly reliable index of minimal proficiency in the home language as well. This was corroborated on a subjective level by the judgement of most examiners that students at around the 70% level seemed to be at least minimally functional in the language.

As to diagnostic analysis, the concept profile seems to represent and to portray visually an accurate view of the degree of control a student has of the structure of the language in question, the peaks and valleys clearly indicating his grammatical strengths and weaknesses. And the profile of an entire group can be graphed by simply totalling and averaging their scores in each concept area and projecting that figure onto the graph. Program and curriculum planning can be effected quite simply for the group by examining, in the handbook itself, the grammatical manifestations corresponding to the concept areas causing the most difficulty. In addition, an overlay of an individual student's profile in English on the profile of his performance in his home language can help us in determining whether the problems he is having in oral communication are linguistic or perhaps developmental in nature.

Preliminary Findings. We can briefly summarize our preliminary findings as follows:

On test revision, there were surprisingly few items that we found needed significant revision.

Two of the visuals in the Listening Comprehension section which provided some difficulty for a great number of students are in the process of being revised.

As to language, only one item in the entire battery of six tests - in the Ilokano Listening Comprehension section - seemed in our opinion to need a change. This is also in the process of being revised.

Reports from our examiners in all languages showed a number of alternate answers in Part III which need to be added to those we already anticipated in the construction of that part of the test.

On the question of dialect variation, the most significant area here is in the Cantonese test. Since a fairly large percentage of the children who took the Cantonese test are from homes where Toi San rather than, or in addition to, Cantonese is used, the extent of differences between Toi San and Cantonese is in the process of being investigated.

If we look at our largest experimental group, the Cantonese, we find that the Cantonese community is, by and large, maintaining the home language, as evidenced by the comparatively large number of children who have functional use of the home language. Of the total Experimental group, 74% were able to complete the three parts of the Cantonese test, and 42% of this group has as good or better control of Cantonese as they do of English.

We're preparing a detailed analysis of their proficiency scores in both Cantonese and English, and we're in the process of doing an item analysis in

each language in order to determine the areas of linguistic difficulties that this group of students is experiencing in each of the two languages. And we'll have similar analyses for the Tagalog, Ilokano, Spanish and Mandarin groups.

It might be of interest that the profile of the Spanish group in Seattle looks quite different from that of the Cantonese group. By and large, the children from Spanish-speaking homes are English dominant, and their overall scores indicate that they are slightly below the monolingual English speakers when they enter school, but fall progressively farther behind through 4th, 5th, and 6th grades.

We are still in the process of analyzing and summarizing our data. When we look at the monolingual English-speaking control group, we find that all of these students scored 70% or higher with scores increasing significantly with age (overall mean: 90%), no doubt reflecting the developmental process of the native language. The mean score for kindergarten students was 80%, with the sixth graders almost topping out at 97%. Future diagnostic analysis will focus on an error analysis of the items missed most frequently, which should supply some indication of the areas of English structure still in the developmental stage at each age/grade level, and at the same time suggest important implications as to curriculum planning for native English-speaking children in the area of English language development.

Reliability and Validity. A reliability study of the English and Spanish tests has just been completed at the University of Idaho (Doyle, 1976) on the Eastern Washington sample which consisted of 209 monolingual English speakers and 291 bilingual students in grades K - 6. The Kuder-Richardson KR-20 formula revealed an overall internal consistency and reliability coefficient of .94 value for the English test and .91 for the Spanish test.

Much work remains to be done, of course, to demonstrate the various kinds of validity of the tests and the extent of each type. However, evidence to date

looks most promising. As to content validity, various linguists, educators, and adult native-speakers have examined the tests and have judged the content to be highly adequate. Developmental changes reflected in age differentiation and sequential patterning of development are apparent in the scores on all six tests, offering initial evidence of construct validity. Significant correlation between oral proficiciency as measured by the MAT-SEA-CAL English test and overall achievement as measured by the SRA Assessment Survey (Science Research Associates) has been established very recently at the .01 confidence level by Talbott (1976) at the University of Idaho, suggesting additional evidence on which to build validity.

Oral Proficiency and Reading Achievement. We also looked into the relationship between oral proficiency and reading achievement scores of our Seattle sample. In order to determine whether there was any significant correlation between spoken language proficiency and reading success, we charted the scores made by a group of children on the MAT-SEA-CAL English test against the scores made by the same children on the MAT reading test (Metropolitan Achievement Test), We were able to obtain MAT reading scores (raw and stanine) on 147 of the total sample of 152 fourth and sixth graders who took the MAT-SEA-CAL English test. This included both the experimental and control groups.

The distribution of the scores reveals that no student who scored below 84% on the oral proficiency test was able to score in the fifth stanine in reading. That is, 90% of those with an oral proficiency score of less than 84% did no better in reading than stanines 1, 2, or 3. In the range between 84% and 90% scores in oral proficiency, we see the emergence of some few five-stanines, but still no sixes. In this range, 70% of the students are still stanine four or below. Only in the above-90% range in oral proficiency do we find appreciable numbers of stanines five through nine.

It would seem, then, that if the fourth and sixth grade student is going to have even the possibility of minimal success in reading (i.e. stanine five), he would have to have sufficient control of spoken English to score 84% or above on the MAT-SEA-CAL. This corresponds to the mean score of the Seattle English-speaking monolingual control group at the kindergarten level. The implications for educational planning and program development contained in these findings would appear to be considerable. This, we believe, is a good thought-provoking notion to leave you with.

REFERENCES

Doyle, Vincent. An Empirical Investigation of the Reliability and Validity of an Oral Proficiency Test. Unpublished dissertation, University of Idaho, Moscow, 1976.

Matluck, Joseph H. and Betty Mace-Matluck. MAT-SEA-CAL Oral Proficiency Tests: English, Spanish, Cantonese, Mandarin, Tagalog, Ilokano, Field Test Edition in 6 volumes, Seattle Public Schools, Seattle, Washington and Center for Applied Linguistics, Arlington, Virginia, 1974.

Talbott, Brian Leigh. The Relationship between Oral Proficiency and Achievement in a Bilingual-Bicultural Elementary School. Unpublished dissertation, University of Idaho, Moscow, 1976.

ABSTRACT

The MAT-SEA-CAL Instruments for Assessing Language Proficiency

Joseph H. Matluck
The University of Texas at Austin

Betty Mace-Matluck
Seattle Public Schools

The Multilingual Test Development Project, recently completed in Seattle, Washington, has yielded a series of comparable oral proficiency tests in six languages designed for use with elementary school children, K - 6. Extensive field-testing of the English, Spanish, Cantonese, Mandarin, Tagalog, and Ilokano tests was conducted with a research sample of several thousand students, both in the metropolitan area of Seattle and in several of the large migrant centers of Eastern Washington and Southern Idaho. This paper describes the project, reports on the preliminary findings, and focuses on the problems encountered and the solutions devised in attempting oral language assessment in a multilingual community. Both the findings and the research design will have broad implications for educational planning and program development by school districts throughout the country and for further research in these areas.